

48

OIPE

## RAW SEQUENCE LISTING

DATE: 10/01/2001

PATENT APPLICATION: US/09/829,275

TIME: 15:17:24

Input Set : A:\4555105.app

Output Set: N:\CRF3\10012001\I829275.raw

3 <110> APPLICANT: WALKER, SUZANNE  
 5 <120> TITLE OF INVENTION: CRYSTALS OF THE ESCHERICHIA COLI MEMBRANE-ASSOCIATED  
 6 GLYCOSYLTRANSFERASE (MURG) PROTEIN, ATOMIC COORDINATES  
 7 AND THREE DIMENSIONAL STRUCTURES THEREOF, ATOMIC  
 8 COORDINATES AND THREE DIMENSIONAL STRUCTURES OF  
 9 BINDING DOMAINS THEREOF, IMAGES THEREOF, AND METHODS  
 10 OF CRYSTALLIZING MURG PROTEINS, MODELS OF UDP-  
 11 GLYCOSYLTRANSFERASES, MURG PROTEINS AND BINDING SITES,  
 12 METHODS OF MAKING MODELS, METHODS OF USING MODELS OF  
 13 MURG, COMPOUNDS THAT BIND, INHIBIT OR STIMULATE MURG  
 14 PROTEINS, AND THERAPEUTIC COMPOSITIONS THEREOF  
 16 <130> FILE REFERENCE: 4555-105  
 18 <140> CURRENT APPLICATION NUMBER: 09/829,275  
 C--> 19 <141> CURRENT FILING DATE: 2001-09-11  
 21 <160> NUMBER OF SEQ ID NOS: 8  
 23 <170> SOFTWARE: PatentIn Ver. 2.1  
 25 <210> SEQ ID NO: 1  
 26 <211> LENGTH: 364  
 27 <212> TYPE: PRT  
 28 <213> ORGANISM: Escherichia coli  
 30 <400> SEQUENCE: 1  
 31 Met Met Ser Gly Gln Gly Lys Arg Leu Met Val Met Ala Gly Gly Thr  
 32 1 5 10 15  
 34 Gly Gly His Val Phe Pro Gly Leu Ala Val Ala His His Leu Met Ala  
 35 20 25 30  
 37 Gln Gly Trp Gln Val Arg Trp Leu Gly Thr Ala Asp Arg Met Glu Ala  
 38 35 40 45  
 40 Asp Leu Val Pro Lys His Gly Ile Glu Ile Asp Phe Ile Arg Ile Ser  
 41 50 55 60  
 43 Gly Leu Arg Gly Lys Gly Ile Lys Ala Leu Ile Ala Ala Pro Leu Arg  
 44 65 70 75 80  
 46 Ile Phe Asn Ala Trp Arg Gln Ala Arg Ala Ile Met Lys Ala Tyr Lys  
 47 85 90 95  
 49 Pro Asp Val Val Leu Gly Met Gly Gly Tyr Val Ser Gly Pro Gly Gly  
 50 100 105 110  
 52 Leu Ala Ala Trp Ser Leu Gly Ile Pro Val Val Leu His Glu Gln Asn  
 53 115 120 125  
 55 Gly Ile Ala Gly Leu Thr Asn Lys Trp Leu Ala Arg Ile Ala Thr Lys  
 56 130 135 140  
 58 Val Met Gln Ala Glu Pro Gly Ala Phe Pro Asn Ala Glu Val Val Gly  
 59 145 150 155 160  
 61 Asn Pro Val Arg Thr Asp Val Leu Ala Leu Pro Leu Pro Gln Gln Arg  
 62 165 170 175  
 64 Leu Ala Gly Arg Glu Gly Pro Val Arg Val Leu Val Val Gly Gly Ser  
 65 180 185 190  
 67 Gln Gly Ala Arg Ile Leu Asn Gln Thr Met Pro Gln Val Ala Ala Lys  
 68 195 200 205

ENTERED

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/829,275

DATE: 10/01/2001  
 TIME: 15:17:24

Input Set : A:\4555105.app  
 Output Set: N:\CRF3\10012001\I829275.raw

```

70 Leu Gly Asp Ser Val Ile Ile Trp His Gln Ser Gly Lys Gly Ser Gln
71      210      215      220
73 Gln Ser Val Glu Gln Ala Tyr Ala Glu Ala Gly Gln Pro Gln His Lys
74 225      230      235      240
76 Val Thr Glu Phe Ile Asp Asp Met Ala Ala Ala Tyr Ala Trp Ala Asp
77      245      250      255
79 Val Val Val Cys Arg Ser Gly Ala Leu Thr Val Ser Glu Ile Ala Ala
80      260      265      270
82 Ala Gly Leu Pro Ala Leu Phe Val Pro Phe Gln His Lys Asp Arg Gln
83      275      280      285
85 Gln Tyr Trp Asn Ala Leu Pro Leu Glu Lys Ala Gly Ala Ala Lys Ile
86      290      295      300
88 Ile Glu Gln Pro Gln Leu Ser Val Asp Ala Val Ala Asn Thr Leu Ala
89 305      310      315      320
91 Gly Trp Ser Arg Glu Thr Leu Leu Thr Met Ala Glu Arg Ala Arg Ala
92      325      330      335
94 Ala Ser Ile Pro Asp Ala Thr Glu Arg Val Ala Asn Glu Val Ser Arg
95      340      345      350
97 Val Ala Arg Ala Leu Glu His His His His His His
98      355      360
101 <210> SEQ ID NO: 2
102 <211> LENGTH: 351
103 <212> TYPE: PRT
104 <213> ORGANISM: Haemophilus influenzae
106 <400> SEQUENCE: 2
107 Met Lys Asn Lys Lys Leu Leu Val Met Ala Gly Gly Thr Gly Gly His
108 1      5      10      15
110 Val Phe Pro Ala Ile Ala Val Ala Gln Thr Leu Gln Lys Gln Glu Trp
111      20      25      30
113 Asp Ile Cys Trp Leu Gly Thr Lys Asp Arg Met Glu Ala Gln Leu Val
114      35      40      45
116 Pro Lys Tyr Gly Ile Pro Ile Arg Phe Ile Gln Ile Ser Gly Leu Arg
117      50      55      60
119 Gly Lys Gly Ile Lys Ala Leu Leu Asn Ala Pro Phe Ala Ile Phe Arg
120 65      70      75      80
122 Ala Val Leu Gln Ala Lys Lys Ile Ile Gln Glu Glu Lys Pro Asp Ala
123      85      90      95
125 Val Leu Gly Met Gly Gly Tyr Val Ser Gly Pro Ala Gly Val Ala Ala
126      100      105      110
128 Lys Leu Cys Gly Val Pro Ile Ile Leu His Glu Gln Asn Ala Ile Ala
129      115      120      125
131 Gly Leu Thr Asn Lys Leu Leu Gly Lys Ile Ala Thr Cys Val Leu Gln
132      130      135      140
134 Ala Phe Pro Thr Ala Phe Pro Met Ala Glu Val Val Gly Asn Pro Val
135 145      150      155      160
137 Arg Glu Asp Leu Phe Glu Met Pro Asn Pro Asp Ile Arg Phe Ser Asp
138      165      170      175
140 Arg Glu Glu Lys Leu Arg Val Leu Val Gly Gly Ser Gln Gly Ala
141      180      185      190

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/829,275

DATE: 10/01/2001

TIME: 15:17:24

Input Set : A:\4555105.app

Output Set: N:\CRF3\10012001\I829275.raw

```

143 Arg Val Leu Asn His Thr Leu Pro Lys Val Val Ala Gln Leu Ala Asp
144      195      200      205
146 Lys Leu Glu Phe Arg His Gln Val Gly Lys Gly Ala Val Glu Glu Val
147      210      215      220
149 Ser Gln Leu Tyr Gly Glu Asn Leu Glu Gln Val Lys Ile Thr Glu Phe
150 225      230      235      240
152 Ile Asp Asn Met Ala Glu Ala Tyr Ala Trp Ala Asp Val Val Ile Cys
153      245      250      255
155 Arg Ser Gly Ala Leu Thr Val Cys Glu Ile Ala Ala Val Gly Ala Ala
156      260      265      270
158 Ala Ile Phe Val Pro Phe Gln His Lys Asp Arg Gln Gln Tyr Leu Asn
159      275      280      285
161 Ala Lys Tyr Leu Ser Asp Val Gly Ala Ala Lys Ile Ile Glu Gln Ala
162      290      295      300
164 Asp Leu Thr Pro Glu Ile Leu Val Asn Tyr Leu Lys Asn Leu Thr Arg
165 305      310      315      320
167 Glu Asn Leu Leu Gln Met Ala Leu Lys Ala Lys Thr Met Ser Met Pro
168      325      330      335
170 Asn Ala Ala Gln Arg Val Ala Glu Val Ile Lys Gln Tyr Ser Asn
171      340      345      350
174 <210> SEQ ID NO: 3
175 <211> LENGTH: 362
176 <212> TYPE: PRT
177 <213> ORGANISM: Enterococcus faecalis
179 <400> SEQUENCE: 3
180 Met Lys Ile Leu Val Thr Gly Gly Gly Thr Gly Gly His Ile Tyr Pro
181 1      5      10      15
183 Ala Leu Ser Phe Val Glu His Val Lys Lys Glu Ala Pro Ala Thr Glu
184      20      25      30
186 Phe Leu Tyr Val Gly Thr Glu Asn Gly Leu Glu Ser Gln Ile Val Pro
187      35      40      45
189 Lys Ala Lys Ile Pro Phe Lys Thr Ile Lys Ile Gln Gly Phe Lys Arg
190      50      55      60
192 Ser Leu Ser Pro Gln Asn Phe Lys Thr Ile Tyr Leu Phe Leu Thr Ser
193 65      70      75      80
195 Ile Asn Lys Ala Lys Lys Ile Ile Arg Glu Phe Gln Pro Asp Val Val
196      85      90      95
198 Ile Gly Thr Gly Gly Tyr Val Ser Gly Ala Val Val Tyr Ala Ala His
199      100      105      110
201 Gln Leu Lys Ile Pro Thr Ile Ile His Glu Gln Asn Ser Ile Pro Gly
202      115      120      125
204 Met Thr Asn Lys Phe Leu Ser Arg Tyr Val Asp Lys Ile Ala Ile Cys
205      130      135      140
207 Phe Pro Asp Val Ala Ser Phe Phe Pro Lys Glu Lys Thr Ile Leu Thr
208 145      150      155      160
210 Gly Asn Pro Arg Gly Gln Glu Val Val Thr Val Glu Lys Ser Ala Ile
211      165      170      175
213 Leu Ser Glu Phe Gly Leu Asp Pro Ala Lys Lys Thr Val Val Leu Phe
214      180      185      190

```

## RAW SEQUENCE LISTING

DATE: 10/01/2001

PATENT APPLICATION: US/09/829,275

TIME: 15:17:24

Input Set : A:\4555105.app

Output Set: N:\CRF3\10012001\I829275.raw

```

216 Gly Gly Ser Arg Gly Ala Leu Lys Ile Asn Gln Ala Phe Glu Gln Ala
217          195          200          205
219 Phe Pro Leu Phe Glu Glu Arg Glu Tyr Gln Val Leu Tyr Ala Ser Gly
220          210          215          220
222 Glu Arg Tyr Tyr Gln Glu Leu Gln Glu Ser Leu Lys Leu Ser Glu Lys
223 225          230          235          240
225 Lys Leu Thr Asn Ile Ser Val Gln Pro Tyr Ile Asp Lys Met Val Glu
226          245          250          255
228 Val Met Ala Asn Thr Asp Leu Met Val Gly Arg Ala Gly Ala Thr Ser
229          260          265          270
231 Ile Ala Glu Phe Thr Ala Leu Gly Leu Pro Ala Ile Leu Ile Pro Ser
232          275          280          285
234 Pro Tyr Val Thr Asn Asp His Gln Thr Lys Asn Ala Gln Ser Leu Val
235          290          295          300
237 Lys Val Gly Ala Val Glu Met Ile Pro Asp Ala Glu Leu Thr Gly Ala
238 305          310          315          320
240 Arg Leu Val Ala Ala Ile Asp Asp Ile Leu Leu Asn Asn Glu Lys Arg
241          325          330          335
243 Gln Gln Met Ala Thr Ala Ser Lys Gly Glu Arg Ile Pro Asp Ala Ser
244          340          345          350
246 Asp Arg Leu Tyr Gln Trp Lys Thr Leu Val
247          355          360
250 <210> SEQ ID NO: 4
251 <211> LENGTH: 360
252 <212> TYPE: PRT
253 <213> ORGANISM: Enterococcus hirae
255 <400> SEQUENCE: 4
256 Met Lys Ile Leu Val Thr Gly Gly Gly Thr Gly Gly His Ile Tyr Pro
257 1          5          10          15
259 Ala Leu Ala Phe Val Asn Tyr Val Lys Thr Lys Glu Pro Asn Thr Glu
260          20          25          30
262 Phe Met Tyr Val Gly Ala Gln Arg Gly Leu Glu Asn Lys Ile Val Pro
263          35          40          45
265 Glu Thr Gly Met Pro Phe Arg Thr Leu Glu Ile Gln Gly Phe Gln Arg
266          50          55          60
268 Lys Leu Ser Leu His Asn Leu Lys Thr Ile Gln Leu Phe Leu Lys Ser
269 65          70          75          80
271 Ile Arg Glu Ala Lys Lys Ile Leu Lys Glu Phe Lys Pro Asp Val Val
272          85          90          95
274 Ile Gly Thr Gly Gly Tyr Val Ser Gly Ala Val Val Tyr Ala Ala Ser
275          100          105          110
277 Lys Leu Ala Ile Pro Thr Ile Ile His Glu Gln Asn Ser Val Pro Gly
278          115          120          125
280 Ile Thr Asn Lys Phe Leu Ser Arg Tyr Val Asp Arg Ile Ala Leu Ser
281          130          135          140
283 Phe Glu Asp Ala Ala Pro Phe Phe Pro Ala Glu Lys Ser Ser Leu Ile
284 145          150          155          160
286 Gly Asn Pro Arg Ala Gln Glu Val Ala Asp Met Asp Lys Ser Lys Ile
287          165          170          175

```

## RAW SEQUENCE LISTING

DATE: 10/01/2001

PATENT APPLICATION: US/09/829,275

TIME: 15:17:24

Input Set : A:\4555105.app

Output Set: N:\CRF3\10012001\I829275.raw

```

289 Leu Ala Thr Tyr Gly Leu Asp Pro Glu Lys Lys Thr Val Leu Ile Phe
290      180      185      190
292 Gly Gly Ser Gln Gly Ala Leu Lys Ile Asn Gln Ala Val Thr Glu Phe
293      195      200      205
295 Leu Met Ser Phe Asp Gln Glu Tyr Gln Val Leu Tyr Ala Ser Gly Glu
296      210      215      220
298 Arg Tyr Tyr Lys Asp Ile Gln Thr Lys Val Pro Ala Cys Ala Asn Val
299 225      230      235      240
301 Ser Ile Gln Pro Tyr Ile Asn Lys Met Ala Glu Val Met Ala Ser Ser
302      245      250      255
304 Asp Leu Leu Val Gly Arg Ala Gly Ala Thr Ser Ile Ala Glu Leu Thr
305      260      265      270
307 Ala Leu Gly Leu Pro Ala Ile Leu Ile Pro Ser Pro Tyr Val Thr Asn
308      275      280      285
310 Asp His Gln Thr Lys Asn Ala Met Ser Leu Val Lys Asn Asn Ala Ala
311      290      295      300
313 Lys Met Ile Lys Asp Asp Glu Leu Asp Gly Arg Ser Leu Lys Gln Ala
314 305      310      315      320
316 Ile Glu Glu Ile Met Thr Asn Asp Gln Leu Gln Lys Gln Met Ser Leu
317      325      330      335
319 Ala Ser Lys Gln Gln Gly Ile Pro Asp Ala Ser Glu Arg Met Tyr Glu
320      340      345      350
322 Leu Val Lys Ser Leu Ile Gln Lys
323      355      360
326 <210> SEQ ID NO: 5
327 <211> LENGTH: 352
328 <212> TYPE: PRT
329 <213> ORGANISM: Streptococcus pneumoniae
331 <400> SEQUENCE: 5
332 Met Lys Lys Ile Val Phe Thr Gly Gly Gly Thr Val Gly His Val Thr
333 1      5      10      15
335 Leu Asn Leu Leu Met Pro Lys Phe Ile Glu Asp Gly Trp Glu Val
336      20      25      30
338 His Tyr Ile Gly Asp Lys Arg Gly Ile Glu His Gln Glu Ile Leu Lys
339      35      40      45
341 Ser Gly Leu Asp Val Thr Phe His Ser Ile Ala Thr Gly Lys Leu Arg
342      50      55      60
344 Arg Tyr Phe Ser Trp Gln Asn Met Leu Asp Val Phe Lys Val Cys Trp
345 65      70      75      80
347 Gly Ile Val Gln Ser Leu Phe Ile Met Leu Arg Leu Arg Pro Gln Thr
348      85      90      95
350 Leu Phe Ser Lys Gly Gly Phe Val Ser Val Pro Pro Val Ile Ala Ala
351      100      105      110
353 Arg Val Ser Gly Val Pro Val Phe Ile His Glu Ser Asp Leu Ser Met
354      115      120      125
356 Gly Leu Ala Asn Lys Ile Ala Tyr Lys Phe Ala Thr Lys Met Tyr Ser
357      130      135      140
359 Thr Phe Glu Gln Ala Ser Ser Leu Ser Lys Val Glu His Val Gly Ala
360 145      150      155      160

```

VERIFICATION SUMMARY  
PATENT APPLICATION: US/09/829,275

DATE: 10/01/2001  
TIME: 15:17:25

Input Set : A:\4555105.app  
Output Set: N:\CRF3\10012001\I829275.raw

L:19 M:271 C: Current Filing Date differs, Replaced Current Filing Date